BGP operations and security draft-jdurand-bgp-security-02.txt

Jerome Durand

Gert Doering

Ivan Pepelnjak

Goals

- Describe BGP security best practices for the Internet
- Synthesis of many existing pieces available (Cymru, RIPE, many IETF docs, some well known pages…)
- Help smaller AS'es build secure and stable BGP networks
- Have consistent recommendations / best practices
- IP version agnostic (IPv4 and IPv6)

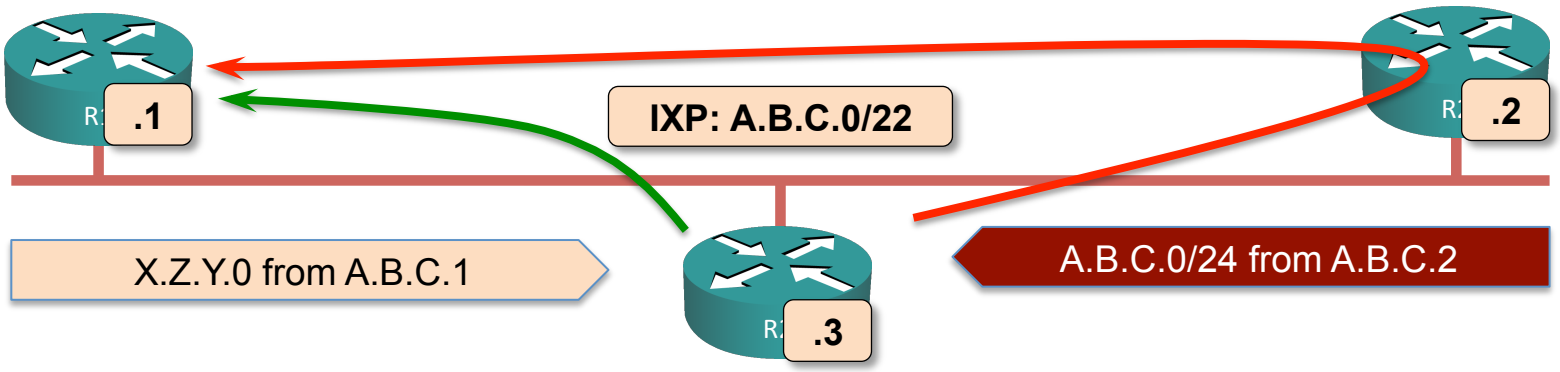
What's covered: the basics

- Control-plane protection (ACL or CoPP)
- BGP session protection (TTL, MD5, TCP-AO)

What's covered: prefix filters

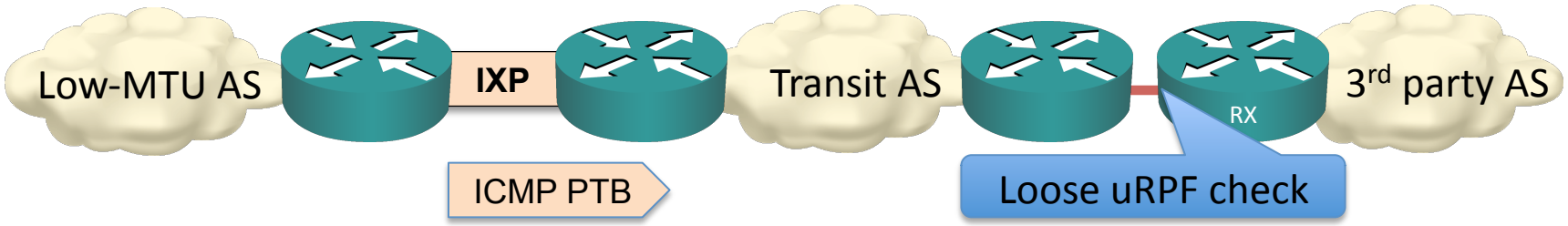
- Default routes
- Special addresses per RFC 5735 (ex: RFC 1918) and IPv6 registry
- Unallocated addresses (IANA and RIR-based)
- RPKI
- Too specific prefixes (descriptive)
- IXP subnets (with examples)

IXP LAN: don't accept more specifics



- More-specific IXP prefixes misdirect traffic and destroy EBGP sessions
- A router MUST NOT accept more specific prefixes for IXP LAN prefix

IXP LAN prefix with pMTUd and uRPF



- ICMP packet sourced from IXP LAN address
- uRPF check might drop ICMP packet ➔ IXP LAN prefix SHOULD be advertised
- Downstream AS might perform strict RIR filter ➔ IXP prefix SHOULD pass RIR filter
- Solution: IXP AS advertises IXP LAN prefix

What's covered: prefix use cases

Use cases

- Full routing networks: filters with peers, upstreams and customers
- Leaf networks

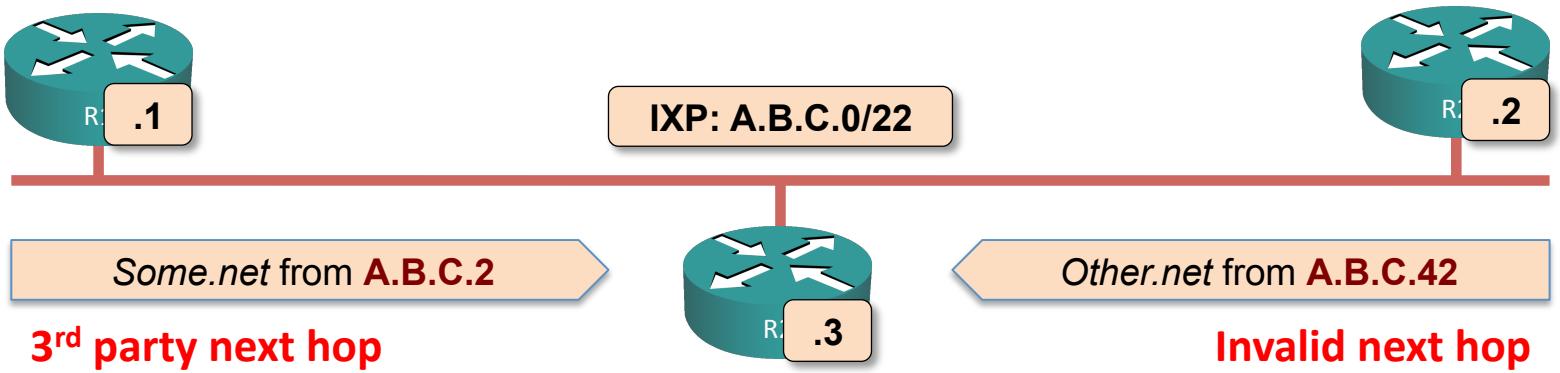
Filters described:

- Inbound and outbound filters
- Loose or strict filters

What's covered: AOB

- BGP Route Flap Dampening (don't)
- Maximum prefixes per peering/BGP neighbor
- AS-path filters (including customer-facing filters)
- Next-hop filters (or next-hop enforcement)
- BGP community scrubbing

BGP next hop filters



- BGP updates can have 3rd party next hop
- Good for optimal traffic flow, bad on IXP LAN
  - Problem#1 – Traffic redirection
  - Problem#2 – Blackholing (invalid next hop)
- **Solution:** change BGP next-hop to peer's IP address with inbound policy

Changes between -02 and -01

- Prefixes were removed from the document and replaced with references to existing registries
- TCP security section now includes TCP-AO
- New section on control-plane protection
- Reworded text about acceptable prefix specificity in former section 4.1.3 to explain this doc does not try to make recommendations
- Remove any reference to anti-spoofing in former section 4.1.4
- Clarification for IXP LAN prefix and pMTUd problem in former section 4.1.5

Changes between -02 and -01

- Replace RIR database with IRR. A definition of IRR is added in former section 4.1.2.2
- 6to4 exception described (only more specifics must be filtered)
- should -> MUST for the prefixes an ISP needs to filter from its customers in former section 4.2.2.1
- Added "plus some headroom to permit growth" in *maximum prefix per peering* section
- Added new section on Next-Hop filtering
- Rewording (ex: Ingress/Egress replaced by Inbound/ Outbound), fixing typos, updated RFC references, and editing changes

Changes between -01 and -00

- Add normative reference for RFC5082 in former section 3.2
  ➔ TTL
- "Non routable" changed in title of former section 4.1.1
  ➔ Prefixes that MUST not be routed by definition
- Correction of typo for IPv4 loopback prefix in former section 4.1.1.1
- Added shared transition space 100.64.0.0/10 in former section 4.1.1.1
- Clarification that 2002::/16 6to4 prefix can cross network boundaries in former section 4.1.1.2
- Rationale of 2000::/3 explained in former section 4.1.1.2
  ➔ In order to build simplified prefix filters
- Added 3FFE::/16 prefix forgotten initially in the simplified list of prefixes that MUST not be routed by definition in former section 4.1.1.2
- Warn that filters for prefixes not allocated by IANA must only be done if regular refresh is guaranteed, with some words about the IPv4 experience, in former section 4.1.2.1

Changes between -01 and -00

- Replace RIR database with IRR. A definition of IRR is added in former section 4.1.2.2
- Remove any reference to anti-spoofing in former section 4.1.4
  ➔ Not anti-spoofing as not dataplane!
- Clarification for IXP LAN prefix and pMTUd problem in former section 4.1.5
  ➔ Using long discussions outcomes on RIPE ML
- "Autonomous filters" typo (instead of Autonomous systems) corrected in the former section 4.2
- Removal of an example for manual address validation in former section 4.2.2.1
- RFC5735 obsoletes RFC3300
- Ingress/Egress replaced by Inbound/Outbound in all the document

Conclusion

- Great feedback received so far!
  - Lot of support and many contributions received
  - Read, Review & Comment!
- Read the document @ https://datatracker.ietf.org/doc/draft-jdurand-bgp-security/
- Discuss on IETF OPSEC WG mailing list @ https://www.ietf.org/mailman/listinfo/opsec

➔Time for IETF OPSEC WG adoption (support)

➔Questions ?